

WO 00/58342

1

PCT/FI00/00249

## SEQUENCE LISTING

&lt;110&gt; Valtion teknillinen tutkimuskeskus

&lt;120&gt; Process for partitioning of molecules

&lt;130&gt; 31805

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 42

&lt;170&gt; PatentIn Ver. 2.2

&lt;210&gt; 1

&lt;211&gt; 428

&lt;212&gt; DNA

<213> *Trichoderma reesei*

&lt;220&gt;

&lt;221&gt; intron

&lt;222&gt; (167)..(236)

&lt;220&gt;

&lt;221&gt; intron

&lt;222&gt; (323)..(386)

&lt;220&gt;

<223> Coding sequence of *hfb1*

&lt;400&gt; 1

```

atgaagttct tcgccatcgc cgctctcttt gccgccgctg ccgttgccca gcctctcgag 60
gaccgcagca acggcaacgg caatgtttgc cctcccgcc tcttcagcaa cccccagtgc 120
tgtgccaccc aagtccctgg cctcatcggc ctgactgca aagtcgtaa gttgagccat 180
aacataagaa tcctcttgac ggaaatatgc cttctcactc ctttaccctt gaacagcctc 240
ccagaacggt tacgacggca ccgacttccg caacgtctgc gccaaaaccg gcgcccagcc 300
tctctgctgc gtggcccccg ttgtaagttg atgcccagc tcaagctcca gtctttggca 360
aaccattctt gacaccaga ctgcaggccg gccaggctct tctgtgccag accgccgtcg 420
gtgcttga                                     428

```

&lt;210&gt; 2

&lt;211&gt; 78

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: PCR 5' primer

&lt;400&gt; 2

```

tcgggcacta cgtgccagta tagcaacgac tactactcgc aatgccttgt tccgcgtggc 60
tctagttctg gaaccgca                                     78

```

<210> 3  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 3  
tcgtacggat cctcaagcac cgacggcggg

30

<210> 4  
<211> 63  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 4  
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagag caacggcaac 60  
ggc 63

<210> 5  
<211> 2211  
<212> DNA  
<213> *Trichoderma reesei*

<220>  
<221> promoter  
<222> (1) .. (2211)  
<223> *cbhl* promoter sequence

<400> 5  
gaattctcac ggtgaatgta ggccttttgt agggtaggaa ttgtcactca agcaccacca 60  
acctccatta cgccctcccc atagagttcc caatcagtga gtcattggcac tgttctcaaa 120  
tagattgggg agaagttgac ttccgcccag agctgaaggt cgcacaaccg catgatatag 180  
ggctcggaac ggcaaaaaag cacgtggctc accgaaaagc aagatgtttg cgatctaaca 240  
tccaggaacc tggatacatc catcatcacg caccgaccact ttgatctgct ggtaaactcg 300  
tattcgccct aaaccgaagt gcgtggtaaa tctacacgtg ggccctttc ggtatactgc 360  
gtgtgtcttc tctaggtgca ttctttcctt cctctagtgt tgaattgttt gtgttgggag 420  
tccgagctgt aactacctct gaatctctgg agaatggtgg actaacgact accgtgcacc 480  
tgcattcatgt atataatagt gatcctgaga aggggggttt ggagcaatgt gggactttga 540  
tgggtcatcaa acaaagaacg aagacgcctc ttttgcaaa ttttgtttcg gctacggtga 600  
agaactggat acttgtttgt tcttctgtgt atttttgtgg caacaagagg ccagagacaa 660  
tctattcaaa caccaagctt gctcttttga gctacaagaa cctgtggggg atatatctag 720  
agttgtgaag tcggtaatcc cgctgtatag taatacagat cgcattctaaa tactccgaag 780  
ctgctgcgaa cccggagaat cgagatgtgc tggaaagctt ctacgcagcg gctaaattag 840  
catgaaaggc tatgagaaat tctggagacg gcttgttgaa tcatggcggt ccattcttcg 900  
acaagcaaa gcttccgtcg cagtagcagg cactcattcc cgaaaaaact cggagattcc 960  
taagtgcgga tggaaaccgga ataataaat aggcataaca ttgagttgcc tcgacggttg 1020  
caatgcagg gtactgagct tggacataac tgttccgtac cccacctctt ctcaaccttt 1080  
ggcgtttccc tgattcagcg tacccgtaca agtcgtaac actattaacc cagactgacc 1140  
ggacgtgttt tgcccttcat ttggagaaat aatgtcattg cgatgtgtaa ttgcctgct 1200  
tgaccgactg gggctgttcg aagcccgaat gtaggattgt tatccgaact ctgctcgtag 1260  
aggcatgttg tgaatctgtg tcgggcagga caccgctcga aggttcacgg caagggaac 1320

```

caccgatagc agtgtctagt agcaacctgt aaagccgcaa tgcagcatca ctggaaaata 1380
caaaccaatg gctaaaagta cataagttaa tgcctaaaga agtcatatac cagcggctaa 1440
taattgtaca atcaagtggc taaacgtacc gtaatttgcc aacgcgttgt ggggttgacg 1500
aagcaacggc aaagccact tcccacgttt gttctttcac tcagtccaat ctgagctggc 1560
gatcccccac ttgggtcgct tgtttgttcc ggtgaagtga aagaagacag aggtaagaat 1620
gtctgactcg gagegttttg catacaacca agggcagtga tgggaagacag tgaaatgttg 1680
acattcaagg agtatttagc cagggatgct tgagtgtatc gtgtaaggag gtttgtctgc 1740
cgatacgacg aatactgtat agtcacttct gatgaagtgg tccatattga aatgtaagtc 1800
ggcactgaac aggcaaaaga ttgagttgaa actgcctaag atctcgggcc ctccgggcttc 1860
ggctttgggt gtacatgttt gtgctccggg caaatgcaaa gtgtggtagg atcgacacac 1920
tgctgccttt accaagcagc tgagggtatg tgataggcaa atgttcaggg gccactgcat 1980
ggtttcgaat agaaagagaa gcttagccaa gaacaatagc cgataaagat agcctcatta 2040
aacgaaatga gctagtaggc aaagtccagc aatgtgtata tataaagggt cgagggtccgt 2100
gcctccctca tgctctcccc atctactcat caactcagat cctccaggag acttgtacac 2160
catcttttga ggcacagaaa cccaatagtc aaccgcggac tgcgcatcat g' 2211

```

&lt;210&gt; 6

&lt;211&gt; 1588

&lt;212&gt; DNA

<213> *Trichoderma reesei*

&lt;220&gt;

<223> *T. reesei* eg11 cDNA

&lt;400&gt; 6

```

ccccctatc ttagtcttcc ttgttgtccc aaaatggcgc cctcagttac actgccgttg 60
accacggcca tcttgcccat tgcccggtc gtcgccgccc agcaaccggg taccagcacc 120
cccagggtcc atcccaagtt gacaacctac aagtgtacaa agtccggggg gtgcgtggcc 180
caggacacct cgggtgtcct tgactggaac taccgctgga tgcacgacgc aaactacaac 240
tcgtgcaccg tcaacggcgg cgtcaacacc acgctctgcc ctgacgaggc gacctgtggc 300
aagaactgct tcatcgaggg cgtcgactac gccgcctcgg gcgtcacgac ctccgggcagc 360
agcctcacca tgaaccagta catgcccagc agctctggcg gctacagcag cgtctctcct 420
cggctgtatc tcttgactc tgacggtgag tacgtgatgc tgaagctcaa cggccaggag 480
ctgagcttcc acgtcgacct ctctgtctcg ccgtgtggag agaacggctc gctctacctg 540
tctcagatgg acgagaacgg gggcgccaac cagtataaca cggccgggtg caactacggg 600
agcggctact gcgatgctca gtgcccgtc cagacatgga ggaacggcac cctcaacact 660
agccaccagg gcttctgctg caacgagatg gatatcctgg agggcaactc gagggcgaat 720
gccttgaccc ctactcttg cacggccacg gcctgcgact ctgccgggtg cggcttcaac 780
ccctatggca gcggctacaa aaactactac ggccccggag ataccgttga cacctccaag 840
accttcacca tcatcaccca gttcaacacg gacaacggct cgcctcggg caaccttgtg 900
agcatcaccc gcaagtacca gcaaaacggc gtcgacatcc ccagcgccca gcccgcgccg 960
gacaccatct cgtcctgccc gtccgcctca gcctacggcg gcctcgccac caggggcaag 1020
gccctgagca gcggcatggg gctcgtgttc agcatttgga acgacaacag ccagtagcatg 1080
aactggctcg acagcggcaa cgcgggcccc tgcagcagca ccgagggcaa cccatccaac 1140
atcttgacca acaaccccaa cacgcacgtc gtcttctcca acatccgctg gggagacatt 1200
gggtctacta cgaactcgac tgcgcccccg ccccgcctg cgtccagcac gacgttttcg 1260
actacacgga ggagctcgac gacttcgagc agcccagact gcacgcagac tactggggg 1320
cagtgcgggt gcattgggta cagcgggtgc aagacgtgca cgtcgggcac tacgtgccag 1380
tatagcaacg actactactc gcaatgcctt tagagcgttg acttgccctc ggtctgtcca 1440
gacgggggca cgatagaatg cgggcacgca gggagctcgt agacattggg cttaatatat 1500
aagacatgct atgttgtatc tacattagca aatgacaaac aaatgaaaaa gaacttatca 1560
agcaaaaaaa aaaaaaaaaa aaaaaaaa 1588

```

<210> 7  
<211> 745  
<212> DNA  
<213> *Trichoderma reesei*  
  
<220>  
<221> terminator  
<222> (1)..(745)  
<223> *T. reesei cbhl* terminator

<400> 7  
ggacctaccc agtctcacta cggccagtgc ggcgggtattg gctacagcgg cccacaggtc 60  
tgcgccagcg gcacaacttg ccaggtcctg aacccttact actctcagt cctgtaaagc 120  
tccgtgcgaa agcctgacgc accggtagat tcttggtgag ccggtatcat gacggcggcg 180  
ggagctacat ggccccgggt gatttatatt ttttgtatct acttctgacc cttttcaa 240  
atcgggtcaa ctcatctttc actggagatg cggcctgctt ggtattgcga tgttgtcagc 300  
ttggcaaatt gtggctttcg aaaacacaaa acgattcctt agtagccatg cattttaaga 360  
taacggaata gaagaaagag gaaattaaaa aaaaaaaaaa aacaaacatc ccgttcataa 420  
cccgtagaat cgccgctctt cgtgtatccc agtaccacgt caaagggtatt catgatcggt 480  
caatgttgat attgttccgc cagtatggct ccaccccat ctccgcgaat ctctcttct 540  
cgaacgcggt agtggctgct gccaatgggt aatgaccata gggagacaaa cagcataata 600  
gcaacagtggt aaattagtgg cgcaataatt gagaacacag tgagaccata gctggcggcc 660  
tggaagcac tgttggagac caacttgctc gttgcgaggc caacttgcac tgctgtcaag 720  
acgatgacaa cgtagccgag gaccc 745

<210> 8  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: annealed primer

<400> 8  
taaccgcggt 10

<210> 9  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: annealed primer

<400> 9  
ctagaccgcg gttaat 16

<210> 10  
<211> 1232  
<212> DNA  
<213> *Trichoderma reesei*

<220>  
<221> promoter  
<222> (1)..(1232)

<223> *T. reesei* *gpd1* promotor

<400> 10

```

gtcgacacga tatacaggcg cggtgatga taatgatgat cgagcatgac ttgatgctgt 60
atgtgacaat attgactgcg aggaaccatc aggtgtgtat ggatggaatc attctgtaac 120
caccaagggtg catgcatcat aaggattctc ctacgtcac caacaacgaa cgatggccat 180
gttagtgaa gacccgtgat ggcaagatag aaccactatt gcactctgctg tccccacgca 240
cagtagctca agtaacgtca aagccgccct cccgtaacct cgcccggtgt tgctccc 300
gattgcctca atcacatagt acctacctat gcattatggg cggcctcaac ccaccccccc 360
agattgagag ctaccttaca tcaatatggc cagcacctct tcggcgatac atactcgcca 420
ccccagccgg cgcgattgtg tgtactaggt aggtctgtac tataccagca ggagagggtgc 480
tgcttggtcaa tctgtctcag ctgttaggtt gtacttgtat ggtacttcta aggtgggtcat 540
gcagttgcta aggtacctag ggagggattc aacgagccct gcttccaatg tccatctgga 600
taggatggcg gctggcgggg ccgaagctgg gaactcgcca acagtcatat gtaatagctc 660
aagttgatga taccgttttg ccagattaga tgcgagaagc agcatgaatg tgcctcatcc 720
gatgcgcgat caccgttttg tcgaaacga ccaagctaag caactaaggt accttaccgt 780
ccactatctc aggttaaccag gtactaccag ctaccctacc tgcgtgctt accctgttta 840
gtgttaatct ttccacctcc ctccctcaatc ttcttttccc tctctctctc tttttttttt 900
cttctctctc ttcttctcca taaccattcc taacaacatc gacattctct cctaataacc 960
agcctcgcaa atcctcagtt tgtatgtacg tacgtactac aatcatcacc acgatcgtcc 1020
gcccagacgat gcggcttctg ttgcctgccc cctcctctca ctctgcccct tgacgagcta 1080
gcccgcgcag gactctctctg cgtcaccaat ttttttccct atttaccct cctccctctc 1140
tccctctcgt ttcttcttaa caacaacca ccacaaaat ctcttggaa gctcacgact 1200
cacgcagctc aattcgcaga tacaatatga ga 1232

```

<210> 11

<211> 1129

<212> DNA

<213> *Trichoderma reesei*

<220>

<221> terminator

<222> (1) ..(1129)

<223> *T. reesei* *gpd1* terminator

<400> 11

```

ggatcccag cattgtctat gaatgcaaac aaaaatagta aataaatagt aattctggcc 60
atgacgaata gagccaatct gctccacttg actatcttgt gactgtatcg tatgtcgaa 120
ccttgactgc ccattcaaac aattgtaaaag gaatatagct acaagttatg tctcacgttt 180
gcgtgcgagc ccgtttgtac gttattttga gaaagcgttg ccatcacatg ctacacgtca 240
cttggtctac gatcatgttt gcgactctcg gtaagaatac acagagtaac gattatctcc 300
atcgcttcta tgattaggtg ctacagacaac acatgggaaa caagataacc atcgcatgca 360
aggtcgattc caatcatgat ctggactggg gtattccatc taagccatag taccctcyag 420
agaaggaatg gtaggacctc tcaggcgctc accatctgtg ctgcaaatcc aagaaaaccc 480
ccaaaagcac ctacctatct acctagagta actgcacgag aaaagaaaag gagcagaaga 540
agaatgatct caagaggccg tgaacgcaga aacacactcc tcccaacttt tcaagttttg 600
aacaacaaaa gaaagatgag gactagaaga tggagtattt ccttctttaga gagctctcgg 660
tgaggtgacc tgtcagggtt taccgcaaac cgtcgggtgg tctatccaat taatcaagtc 720
ccgcgcctcg cctcttctct cctgtccttt catagaatcc cgtctccttg ttgcttgatc 780
gaagcggggg tatcgacgac accaaagatc ttgtcttggg gacttatcaa tctttgggtg 840
atcaaacagc ccccgagtga tcagatccgt aaaagaagaa gaagagtacg atttaaccag 900
accgaggaac aataaagcga gtaaataaca tcaaaaataag agtctcgttg aaaattactt 960
gttcctcaat caatcccaac ccccctaaaa gcccttcccc ccattgggata tcccggcagt 1020
aggagagaga tatttccact accgctcacc accaagttag gcttgccgag agaagaggat 1080
gaatcagaag tgacaacaac ggggttagca catgggatat cggcgcgcc 1129

```

<210> 12  
<211> 5733  
<212> DNA  
<213> *Aspergillus nidulans*  
  
<220>  
<223> (1-5733) Sequence of plasmid pAN52-1

<220>  
<221> promoter  
<222> (1)..(2129)  
<223> *A. nidulans* *gpdA* promoter

<220>  
<221> gene  
<222> (2130)..(2304)  
<223> *A. nidulans* *gpdA* gene

<220>  
<221> terminator  
<222> (2305)..(3071)  
<223> *A. nidulans* *trpC* terminator

<220>  
<221> misc\_feature  
<222> (3072)..(5726)  
<223> pUC18 from Sali I to EcoRI

<400> 12  
caattccctt gtatctctac acacaggctc aaatcaataa gaagaacggt tcgtcttttt 60  
cgtttatata ttgcatcgct ccaaagctat tggcgggata ttctgtttgc agttggctga 120  
cttgaagtaa tctctgcaga tctttcgaca ctgaaatacg tcgagcctgc tccgcttgga 180  
agcggcgagg agcctcgctc tgtcacaaact accaacaagg agtacgataa gggccagttc 240  
cgccagctca ttaagagcca gttcatgggc gttggcatga tggccgctcat gcatctgtac 300  
ttcaagtaca ccaacgctct tctgatccag tcgatcatcc gctgaaggcg ctttcgaatc 360  
tggttaagat ccacgtcttc gggaaaggcg cgactgggtga cctccagcgt ccttttaagg 420  
ctgccaacag ctttctcagc cagggccagc ccaagaccga caaggcctcc ctccagaacg 480  
cggagaagaa ctggagggggt ggtgtcaagg aggagtaagc tccttattga agtcggagga 540  
cggagcgggt tcaagaggat attcttcgac tctgtattat agataagatg atgaggaatt 600  
ggaggttagc tagcttcatt tggatttgct ttccaggctg agactctagc ttggagcata 660  
gagggtcctt tggctttcaa tattctcaag tatctcgagt ttgaacttat tccctgtgaa 720  
ccttttattc accaatgagc attggaatga acatgaatct gaggactgca atcgccatga 780  
ggttttcgaa atacatccgg atgtcgaagg ctgggggcac ctgctgttgg tgaatttaga 840  
acgtggcact attgatcatc cgatagctct gcaaaggcg tgcacaaatg caagtcaaac 900  
gttgctagca gttccagggt gaatgttatg atgagcattg tattaatatg ggagatatag 960  
catgatctct agttagctca ccacaaaagt cagacggcgt aacccaaaagt cacacaacac 1020  
aagctgtaag gatttcggca cggctacgga agacggagaa gccaccttca gtggactcga 1080  
gtaccattta attctatttg tgtttgatcg agacctata cagccctac aacgaccatc 1140  
aaagtcgtat agctaccagt gagggaagtgg actcaaatcg acttcagcaa catctcctgg 1200  
ataaacttta agcctaaact atacagaata agataggtgg agagcttata ccgagctccc 1260  
aaatctgtcc agatcatggg tgaccgggtgc ctggatcttc ctatagaatc atccttattc 1320  
gttgacctag ctgattctgg agtgaccagg agggctatga cttgagccta aaatccgccg 1380  
cctccaccat ttgtagaaaa atgtgacgaa ctctgtagct ctgtacagt accgggtgact 1440  
ctttctggca tgcggagaga cggacggagc cagagagaag ggctgagtaa taagccactg 1500  
gccagacagc tctggcgggt ctgaggtgca gtggatgatt attaatccgg gaccggccgc 1560  
ccctccgccc cgaagtggaa aggtctgggt gccctcgtt gaccgaagaat ctattgcatc 1620  
atcggagaat atggagcttc atcgaatcac cggcagtaag cgaaggagaa tgtgaagcca 1680  
ggggtgtata gccgtcggcg aaatagcatg ccattaacct aggtacagaa gtccaattgc 1740

ttccgatctg gtaaaagatt caccgagatag taccttctcc gaagtaggta gagcgagtag 1800  
 cgggcgcgta agctccctaa ttggcccatc cggcatctgt agggcgctcca aatatcgtag 1860  
 ctctctctgt ttgcccgggtg tatgaaaccg gaaaggccgc tcaggagctg gccagcggcg 1920  
 cagaccggga acacaagctg gcagtcgacc catccggtag tctgactcg acctgctgag 1980  
 gtccctcagt ccctggtagg cagctttgcc cgtctgtcc gccgggtgtg tcggcggggg 2040  
 tgacaaggtc gttgctcag tccaacatct gttgccatat ttctctgtc tccccaccag 2100  
 ctgctctttt cttttctctt tcttttccca tcttcagtat attcatcttc ccatccaaga 2160  
 acctttattt cccctaagta agtactttgc tacatccata ctccatcctt cccatccctt 2220  
 attcctttga acctttcagt tcgagctttc ccacttcac gcagcttgac taacagctac 2280  
 cccgcttgag cagacatcac catggatcca cttaacgtta ctgaaatcat caaacagctt 2340  
 gacgaatctg gatataagat cgttgggtgc gatgtcagct ccggaggtga gacaaatggt 2400  
 gttcaggatc tcgataagat acgttcattt gtccaagcag caaagagtgct cttctagtga 2460  
 tttaatagct ccatgtcaac aagaataaaa cgcgttttcg ggtttacctc tccagatac 2520  
 agctcatctg caatgcatta atgcattgac tgcaacctag taacgccttn cagggtcccg 2580  
 cgaagagaag aatagcttag cagagctatt ttcattttcg ggagacgaga tcaagcagat 2640  
 caacggctgt caagagacct acgagactga ggaatccgct ctgggtcca cgcgactata 2700  
 tatttctctc taattgtact ttgacatgct cctcttctt actctgatag cttgactatg 2760  
 aaaatctcgt caccagcnc cgggttcgca aagataattg catgtttctt ccttgaactc 2820  
 tcaagcctac aggacacaca ttcacgttag gtataaacct cgaatcant tcctactaag 2880  
 atggatatac atagtaacca tgcatgggtg cctagtgaat gctccgtaac acccaatac 2940  
 cggccgaaa cttttttaca actctcctat gagtctgtta ccagaatgc acaggtacac 3000  
 ttgttttagg gtaatccttc tttctagaag tcctcgtgta ctgtgtaagc gccactcca 3060  
 catctccact cgacctgcag gcatgcaagc ttggcactgg ccgtcgtttt acaacgtcgt 3120  
 gactgggaaa accctggcgt taccacactt aatcgcttg cagcacatcc cctttcggc 3180  
 agctggcgtg atagcgaaga ggcccgacc gatcgccctt cccaacagtt gcgcagcctg 3240  
 aatggcgaat ggccgctgat gcggtatttt ctcttacgc atctgtgcgg tatttcacac 3300  
 cgcataatggt gcaactctcag tacaatctgc tctgatgcg catagttaag ccagccccga 3360  
 caccgcgcaa caccgcgtga cgcgcctga cgggcttgct tgctcccgcc atccgcttac 3420  
 agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg 3480  
 aaacgcgca gacgaaagg cctcgtgata cgcctatctt tatagggtta tgctatgata 3540  
 ataattggtt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacctctatt 3600  
 tgtttatttt tctaataaca ttcaaatatg tatccgctca tgagacaata acctgataa 3660  
 atgttcaaat aatatgaaa aagggaagat atgagtattc aactttccg tgtccgctt 3720  
 attcctttt ttgcggcatt ttgccttctt gtttttgctc acccagaaac gctggtgaaa 3780  
 gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 3840  
 agcgtgaaga tccttgagag ttttcgcccc gaagaacgtt tccaatgat gagactttt 3900  
 aaagtctcgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcgg 3960  
 cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcgt 4020  
 cttacggatg gcatgacagt aagagaatta tgcagtgcgt ccataaccat gagtataac 4080  
 actgcggcca acttacttct gacaacgac ggaggaccga aggagctaac cgctttttt 4140  
 cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagc 4200  
 ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaa 4260  
 ctattaactg gcgaactact tactctagct tcccggaac aattaataga ctggatggag 4320  
 gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttatgtc 4380  
 gataaatctg gagccgggtg gcgtgggtct cgcggtatca ttgcagcact ggggccagat 4440  
 ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 4500  
 cgaaatagac agatcgctga gatagggtgc tcaactgatta agcattggta actgtcagac 4560  
 caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 4620  
 taggtgaaga tcctttttga taaatctcat accaaaatcc cttaacgtga gttttcgttc 4680  
 cactgagcgt cagacccgt agaaaagatc aaaggatctt cttagatcc tttttttctg 4740  
 cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcgggtgtt ttgtttgccc 4800  
 gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 4860  
 aatactgtcc ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg 4920  
 cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 4980  
 tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 5040  
 acgggggggt cgtgcacaca gcccgcttg gagcgaacga cctacaccga actgagatac 5100  
 ctacagcgtg agctatgaga aagcggcacg cttcccgaag ggagaaaggc ggacaggtat 5160  
 ccggtaaagc gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaacgcc 5220

tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 5280  
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt ttacgggttc 5340  
ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg 5400  
gataaccgta ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag 5460  
cgcagcgagt cagtgagega ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc 5520  
gcgcgttggc cgattcatta atgcagctgg cagcacaggt ttcccgaact gaaagcgggc 5580  
agtgagegca acgcaattaa tgtgagttag ctactcatt aggcacccca ggctttacac 5640  
tttatgcttc cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga 5700  
aacagctatg accatgatta cgaattgcgg ccg 5733

<210> 13

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 5' primer

<400> 13

gtcaaccgcg gactgcgcat catgaagttc ttgccatc 39

<210> 14

<211> 66

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 5' primer

<400> 14

tctagcaagc ttggctctag ttctggaacc gcaccaggcg gcagcaacgg caacggcaat 60  
gtttgc 66

<210> 15

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 3' primer

<400> 15

tcgtacaagc tttcaagcac cgacggcggt 30

<210> 16

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 5' primer

<400> 16

tctagctcta gaagcaacgg caacggcaat gtt 33



<210> 17  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 17  
tgctagtcga cctgctagca gcaccgacgg cggctctg

37

<210> 18  
<211> 4614  
<212> DNA  
<213> *Saccharomyces cerevisiae*

<220>  
<223> *S. cerevisiae* FLO1 coding sequence

<400> 18  
atgacaatgc ctcatcgcta tatgtttttg gcagtcctta cacttctggc actaactagt 60  
gtggcctcag gagccacaga ggcgtgctta ccagcaggcc agaggaaaag tgggatgaat 120  
ataaattttt accagtatct attgaaagat tcctccacat attcgaatgc agcatatatg 180  
gcttatggat atgcctcaaa aaccaaacta ggttctgtcg gaggacaaac tgatatctcg 240  
attgattata atattccctg tgttagttca tcaggcacat ttccttgtcc tcaagaagat 300  
tcctatggaa actggggatg caaaggaatg ggtgcttgtt ctaatagtca aggaattgca 360  
tactggagta ctgatttatt tggtttctat actaccccaa caaacgtaac cctagaaatg 420  
acaggttatt ttttaccacc acagacgggt tcttacacat tcaagtttgc tacagttgac 480  
gactctgcaa ttctatcagt aggtgggtgca accgcgttca actgttgtgc tcaacagcaa 540  
ccgccgatca catcaacgaa ctttaccatt gacggatca agccatgggg tggagtttg 600  
ccaccttaata tcgaaggaa cgtctatatg tacgctggct actattatcc aatgaaggtt 660  
gtttactcga acgctgttcc ttgggggtaca cttccaatta gtgtgacact tccagatggg 720  
accactgtaa gtgatgactt cgaagggtac gtctattcct ttgacgatga cctaagtcaa 780  
tctaactgta ctgtccctga cccttcaaat tatgctgtca gtaccactac aactacaacg 840  
gaaccatgga ccggtacttt cacttctaca tctactgaaa tgaccaccgt caccggtacc 900  
aacggcggtt caactgacga aaccgtcatt gtcacagaa ctccaacaac tgctagcacc 960  
atcataacta caactgagcc atggaacagc acttttacct ctacttctac cgaattgacc 1020  
acagtcactg gcaccaatgg tgtacgaact gacgaaacca tcattgtaat cagaacacca 1080  
acaacagcca ctactgccat aactacaact gagccatgga acagcacttt tacctctact 1140  
tctaccgaat tgaccaacagt caccgggtacc aatgggttgc caactgatga gaccatcatt 1200  
gtcatcagaa caccaacaac agccactact gccatgacta caactcagcc atggaacgac 1260  
acttttacct ctacatccac tgaaatgacc accgtcaccg gtaccaacgg ttgccaact 1320  
gatgaaacca tcattgtcat cagaacacca acaacagcca ctactgctat gactacaact 1380  
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccaccgt caccgggtacc 1440  
aacgggttgc caactgatga aaccatcatt gtcacagaa caccaacaac agccactact 1500  
gccatgacta caactcagcc atggaacgac acttttacct ctacatccac tgaaatgacc 1560  
accgtcaccg gtaccaatgg ttgccaact gatgagacca tcattgtcat cagaacacca 1620  
acaacagcca ctactgccat gactacaact cagccatgga acgacacttt tacctctaca 1680  
tccactgaaa tgaccaccgt caccgggtacc aacgggttgc caactgatga aaccatcatt 1740  
gtcatcagaa caccaacaac agccactact gccataacta caactgagcc atggaacagc 1800  
acttttacct ctacttctac cgaattgacc acagtcaccg gtaccaatgg ttgccaact 1860  
gatgagacca tcattgtcat cagaacacca acaacagcca ctactgccat gactacaact 1920  
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccaccgt caccgggtacc 1980  
aacgggttgc caactgatga aaccatcatt gtcacagaa caccaacaac agccactact 2040  
gccatgacta caactcagcc atggaacgac acttttacct ctacatccac tgaaatgacc 2100  
accgtcaccg gtaccaacgg ttgccaact gatgagacca tcattgtcat cagaacacca 2160

acaacagcca ctactgccat gactacaact cagccatgga acgacacttt tacctctaca 2220  
 tccactgaaa tgaccaccgt caccgggtacc aacggcggtc caactgacga aaccgtcatt 2280  
 gtcacagaaa ctccaactag tgaagggtcta atcagcacca ccaactgaacc atggactggg 2340  
 actttcacct ctacatccac tgagatgacc accgtcacgg gtactaacgg tcaaccaact 2400  
 gacgaaacgg tgattgttat cagaactcca accagtgaag gtttggttac aaccaccact 2460  
 gaaccatgga ctgggtacttt tacttctaca tctactgaaa tgaccaccat tactggaacc 2520  
 aacggcggtc caactgacga aaccgtcatt gtcacagaaa ctccaaccag tgaagggtcta 2580  
 atcagcacca ccaactgaacc atggactggg acttttactt ctacatctac tgaatgacc 2640  
 accattactg gaaccaatgg tcaaccaact gacgaaacgg ttattgttat cagaactcca 2700  
 actagtgaag gtctaactag caccaccact gaaccatgga ctgggtacttt cacttctaca 2760  
 tctactgaaa tgaccaccgt caccgggtacc aacggcggtc caactgacga aaccgtcatt 2820  
 gtcacagaaa ctccaaccag tgaagggtcta atcagcacca ccaactgaacc atggactggg 2880  
 actttcactt cgacttccac tgaggttacc accatcactg gaaccaacgg tcaaccaact 2940  
 gacgaaactg tgattgttat cagaactcca accagtgaag gtctaactag caccaccact 3000  
 gaaccatgga ctgggtacttt cacttctaca tctgtgaaa tgaccaccgt caccgggtact 3060  
 aacgggtcac caactgacga aaccgtgatt gttatcagaa ctccaaccag tgaagggttg 3120  
 gttacaacca ccaactgaacc atggactggg acttttactt cgacttccac tgaatgtct 3180  
 actgtcactg gaaccaatgg cttgccaaact gatgaaactg tcattgttgt caaaactcca 3240  
 actactgcca tctcatccag tttgtcatca tcatcttcag gacaaatcac cagctctatc 3300  
 acgtcttcgc gtccaattat taccctattc tatcctagca atggaacttc tgtgatttct 3360  
 tctcagtaa tttcttctc agtcaattct tctctattca cttcttctcc agtcatctct 3420  
 tctcagtaa tttcttctc tacaacaacc tccacttcta tattttctga atcatctaaa 3480  
 tcatccgtca ttccaaccag tagttccacc tctggttctt ctgagagcga aacgagttca 3540  
 gctggttctg tctcttctc ctcttttacc tcttctgaat catcaaaatc tctacatat 3600  
 tcttctctat cattaccact tgttaccagt gcgacaacaa gccagggaac tgccttctca 3660  
 ttaccactg ctaccactac aaaaacgagc gaacaaacca ctttggttac cgtgacatcc 3720  
 tgcgagtctc atgtgtgcac tgaatccatc tccctgcga ttgtttccac agtactgtt 3780  
 actgttagcg gcgtcacaac agagtatacc acatggtgcc ctatttctac tacagagaca 3840  
 acaaagcaaa ccaaaggagc aacagagcaa accacagaaa caacaaaaca aaccacggta 3900  
 gttacaattt cttctgtgta atctgacgta tgctctaaga ctgcttctcc agccattgta 3960  
 tctacaagca ctgctactat taacggcggt actacagaat acacaacatg gtgtcctatt 4020  
 tccaccacag aatcgaggca acaacaacg ctagtactg ttacttctct cgaatctggg 4080  
 gtgtgttccg aaactgctc acctgccatt gtttcgacgg ccacggctac tgtgaatgat 4140  
 gtgtgtacgg tctatctac atggaggcca cagactgcga atgaagagtc tgtcagctct 4200  
 aaaatgaaca gtgctaccgg tgagacaaca accaatactt tagctgctga aacgactacc 4260  
 aatactgtag ctgctgagac gattaccaat actggagctg ctgagacgaa aacagtagtc 4320  
 acctcttcgc tttcaagatc taatcacgct gaaacacaga cggcttccgc gaccgatgtg 4380  
 attggtcaca gcagtagtgt tgtttctgta tccgaaactg gcaacaccaa gactctaaca 4440  
 agttccgggt tgagtactat gtcgcaacag cctcgtagca caccagcaag cagcatggta 4500  
 ggatatagta cagcttcttt agaaatttca acgtatgctg gcagtgcgaa cagcttactg 4560  
 gccggtagtg gtttaagtgt cttcattgct tccttattgc tggcaattat ttaa 4614

<210> 19

<211> 5857

<212> DNA

<213> *Saccharomyces cerevisiae* and *E. coli*

<220>

<221> promoter

<222> (1) .. (452)

<223> *S. cerevisiae* GAL1 promoter

<220>

<223> (476-495) *E. coli* T7 promoter/priming site

<220>

<223> (502-601) *E. coli* multiple cloning site

<220>  
<223> (609-857) *S. cerevisiae* CYC1 transcription terminator

<220>  
<223> (1039-1712) *E. coli* pMB1 (pUC-derived) origin

<220>  
<221> gene  
<222> (1857)..(2717)  
<223> *E. coli* ampicillin resistance gene

<220>  
<221> gene  
<222> (2735)..(3842)  
<223> *S. cerevisiae* URA3 gene

<220>  
<223> (3846-5317) *S. cerevisiae* 2 micron origin

<220>  
<223> (5385-5840) *E. coli* f1 origin

<220>  
<223> (1-5857) Sequence of pYES2

<400> 19  
acggattaga agccgcccag cgggtgacag ccctccgaag gaagactctc ctccgtgcgt 60  
cctcgctctc accggctcgcg ttcttgaaac gcagatgtgc ctgcgcccgc actgctccga 120  
acaataaaga ttctacaata ctgcttttta tggttatgaa gaggaaaaat tggcagtaac 180  
ctggcccccac aaaccttcaa atgaacgaat caaattaaca accataggat gataatgcca 240  
ttagtttttt agccttattt ctggggtaat taatcagcga agcgatgatt ttgatctat 300  
taacagatat ataaatgcaa aaactgcatt aaccacttta actaatactt tcaacatttt 360  
cggtttgtat tacttcttat tcaaatgtaa taaaagtatc acaaaaaaat tgtaataata 420  
cctctatact ttaacgtcaa ggagaaaaaa ccccggtatcg gactactagc agctgtaata 480  
cgactcacta tagggaatat taagcttggg accgagctcg gatccactag taacggccgc 540  
cagtggtgctg gaattctgca gatattccatc acactggcgg ccgctcagac atgcatctag 600  
agggccgcag catgtaatta gttatgtcac gcttacattc acgcccctcc cccacattcc 660  
ctctaaccga aaaggaagga gttagacaac ctgaagtcta ggtccctatt tattttttta 720  
tagttatggt agtattaaga acgttattta tatttcaaat ttttcttttt tttctgtaca 780  
gacgcgtgta cgcattgaac attatactga aaaccttgct tgagaagggt ttgggacgct 840  
cgaaggcttt aatttgcggc cctgcattaa tgaatcggcc aacgcgcggg gagaggcgg 900  
ttgcgtattg ggcgctcttc cgttctctcg ctactgact cgtgcgctc ggtcgttcgg 960  
ctgcggcgag cggtatcagc tcaactcaaag gcgtaatac gggtatccac agaatacagg 1020  
gataaacgcag gaaagaacat gtgagcaaaa ggccagcaaa agcccaggaa ccgtaaaaaa 1080  
gccgcgttgc tggcgttttt ccataggctc cgcctccctg acgagcatca caaaaatcga 1140  
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 1200  
ggaagctccc tcgtgcgctc tctgttcccg accctgcgcg ttaccggata cctgtccgcc 1260  
tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 1320  
gtgtaggctg ttgcgtccaa gctgggctgt gtgcacgaac ccccgttca gcccgaccgc 1380  
tgcgccttat ccggaacta tcgtcttgag tccaaaccgg taagacacga cttatcgcca 1440  
ctggcagcag cactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 1500  
ttcttgaaat ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 1560  
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc 1620  
accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 1680  
ttctaagaag atcctttgat cttttctacg gggtctgacg ctcatgggaa cgaaaactca 1740  
cgtaaggga ttttggcat gagattatca aaaaggatct tcacctagat ccttttaaat 1800  
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac 1860  
caatgcttaa tcagtgggc acctatctca gcgatctgtc tatttcgttc atccatagtt 1920

gcctgactcc cgcgcgtgta gataactacg atacgggagc gcttaccatc tggccccagt 1980  
gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 2040  
ccagccggaa gggccgagcg cagaagtggc cctgcaactt tatccgcctc cattcagctc 2100  
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgct 2160  
gttggcattg ctacaggcat cgtgggtgca ctctcgctcg ttggatggc ttcatccagc 2220  
tccgggtccc aacgatcaag gcgagttaca tgatcccca tgttggtgca aaaagcgggt 2280  
agctccctcg gtcctccgat cgttgtcaga agtaagtgg ccgcagtggt atcactcatg 2340  
gttatggcag cactgcataa ttctcttact gtcagccat ccgtaagatg cttttctgtg 2400  
actgggtgagt actcaaccaa gtcattctga gaatagtga tgcggcgacc gagggtgctc 2460  
tgcccgcgct caatacggga taatagtga tcacatagca gaactttaaa agtgctcatc 2520  
attggaaaaa gttcttcggg gcgaaaaact tcaaggatct taccgctggt gagatccagt 2580  
tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt 2640  
tctgggtgag caaaaacagg aaggcaaat gccgcaaaaa agggaataag ggcgacacgg 2700  
aaatgttgaa tactcatact ctctcttttt caatgggtaa taactgatat aattaaattg 2760  
aagctctaatt ttgtgagttt agtatacatg catttactta taatacagtt ttttagtttt 2820  
gctggccgca tcttctcaaa tatgcttccc agcctgcttt tctgtaacgt tcacctcta 2880  
ccttagcatc ccttcccttt gcaaatagtc ctcttcaac aataataatg tcagatcctg 2940  
tagagaccac atcatccacg gttctatact gttgacccaa tgcgtctccc ttgtcatcta 3000  
aaccacacac ggggtgcata atcaaccaat cgtaaccttc atctcttcca cccatgtctc 3060  
tttgagcaat aaagccgata acaaaatctt tgcgtctctt cgcaatgtca acagtacctc 3120  
tagtatattc tccagtagat agggagccct tgcagacaaa ttctgctaac atcaaaaggc 3180  
ctctaggttc ctttgttact tcttctgcgc cctgcttcaa accgctaaca atacctgygc 3240  
ccaccacacc gtgtgcattc gtaatgtctg cccattctgc tattctgtat acaccgcgag 3300  
agtactgcaa tttgactgta ttaccaatgt cagcaaattt tctgtcttcg aagagtaaaa 3360  
aattgtactt ggcggataat gcctttagcg gcttaactgt gccctccatg gaaaaatcag 3420  
tcaagataat cacatgtggt ttttagtaaac aaattttggg acctaatgct tcaactaact 3480  
ccagtaattc cttggtggta cgaacatcca atgaagcaca caagtttgtt tgcttttctg 3540  
gcatgatatt aaatagcttg gcagcaacag gactaggatg agtagcagca cgttccttat 3600  
atgtagcttt cgacatgatt tatcttcggt tcttgcaggt ttttgttctg tgcagttggg 3660  
ttaagaatac tgggcaattt catgtttctt caacactaca tatgcgtata tataccaatc 3720  
taagtctgtg ctctctctt cgttcttctt tctgttcgga gattaccgaa tcaaaaaaat 3780  
ttcaaagaaa ccgaaatcaa aaaaaaagaa aaaaaaaaaa tgatgaattg aattgaaaag 3840  
ctagcttata gatgataagc tgtcaaagat gagaattaat tccacggact atagactata 3900  
ctagatactc cgtctactgt acgatacact tccgctcagg tccttgtctt ttaacgaggc 3960  
cttaccactc ttttgttact ctattgatcc agctcagcaa aggcagtggt atctaagatt 4020  
ctatcttcgc gatgtagtaa aactagctag accgagaaaag agactagaaa tgcaaaaggc 4080  
acttctacaa tggctgccat cattattatc cgatgtgacg ctgcagcttc tcaatgatat 4140  
tcgaatacgc tttgaggaga tacagcctaa tatccgacaa actgttttac agatttacga 4200  
tcgtacttgt taccatcat tgaattttga acatccgaac ctgggagttt tccctgaaac 4260  
agatagtata tttgaacctg tataataata tatagtctag cgctttacgg aagacaatgt 4320  
atgtatttcg gttctcggag aaactattgc atctattgca taggtaatct tgcacgtcgc 4380  
atccccgggt cattttctgc gtttccatct tgcacttcaa tagcatatct ttgttaacga 4440  
agcatctgtg cttcattttg tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac 4500  
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tattttacca 4560  
acgaagaatc tgtgtctcat ttttgtaaaa caaaaatgca acgcgacgag agcgctaatt 4620  
tttcaaacaa agaactgtag ctgcattttt acagaacaga aatgcaacgc gagagcgcta 4680  
ttttaccaac aaagaatcta tacttctttt ttgttctaca aaaatgcac ccgagagcgc 4740  
tatttttcta acaaagcatc tttagattact ttttttctcc tttgtgcgct ctataatgca 4800  
gtctcttgat aactttttgc actgtaggtc cgtaagggtt agaagaaggc tactttgggt 4860  
tctattttct cttccataaa aaaagcctga ctccacttcc cgcgtttact gattactagc 4920  
gaagctgcgg gtgcattttt tcaagataaa ggcattcccc atttatattct ataccgatgt 4980  
ggattgcgca tactttgtga acagaaagtg atagcgttga tgattcttca ttggtcagaa 5040  
aattatgaac ggtttcttct attttgtctc tatatactac gtataggaaa tgtttacatt 5100  
ttcgtattgt tttcgattca ctctatgaat agttcttact acaatttttt tgtctaaaga 5160  
gtaatactag agataaacat aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag 5220  
cgaaagggtg atgggtagggt tatataggga tatagcacag agatatatag caaagagata 5280  
cttttgagca atgtttgtgg aagcggattt cgcaatggga agctccaccc cggttgataa 5340  
tcagaaaagc cccaaaaaca ggaagattgt ataagcaaat atttaaatg taaacgttaa 5400

```

tattttgtta aaattcgcgt taaatttttg ttaaatacagc tcatttttta acgaatagcc 5460
cgaaatcggc aaaatccctt ataaatcaaa agaataagacc gagatagggt tgagtgttgt 5520
tccagtttcc aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa 5580
aagggtctat cagggcgatg gccactacg tgaaccatca ccctaataca gttttttggg 5640
gtcgagggtgc cgtaaagcag taaatcggaa gggtaaacgg atgcccccat ttagagcttg 5700
acgggggaaag ccggcgaaacg tggcgagaaa ggaagggaag aaagcgaaaag gagcgggggc 5760
tagggcggtg ggaagtgtag gggtcacgct gggcgtaacc accacacccg ccgcgcttaa 5820
tggggcgcta cagggcgcgt ggggatgac cactagt 5857

```

<210> 20  
 <211> 403  
 <212> DNA  
 <213> *Trichoderma reesei*

<220>  
 <223> (1-403) *T. reesei* hfb2 coding sequence

<220>  
 <221> intron  
 <222> (131) .. (200)

<220>  
 <221> intron  
 <222> (287) .. (358)

<400> 20  
 atgcagttct tgcgcgtgc cctcttcgcc accagcgccc tggctgctgt ctgccctacc 60  
 ggccctcttct ccaaccctct gtgctgtgcc accaacgtcc tcgacctcat tggcggtgac 120  
 tgcaagaccc gtatgttgaa ttccaatctc tgggcatect gacattggac gatacagttg 180  
 acttacacga tgctttacag ctaccatcgc cgtcgacact ggcgccatct tccaggctca 240  
 ctgtgccagc aagggtcca agcctctttg ctgctgtgct cccgtggtaa gtagtgctcg 300  
 caatggcaaa gaagtaaaaa gacatttggg cctgggatcg ctaactcttg atatcaaggc 360  
 cgaccaggct ctctgtgccc agaaggccat cggcaccttc taa 403

<210> 21  
 <211> 59  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PCR 5' primer

<400> 21  
 cggaggagct cgacgacttc gagcagcccg agctgcacgc aggtgtgtctg ccctaccgg 59

<210> 22  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PCR 3' primer

<400> 22  
 tcattggatc cttagaaggt gccgatggc

<210> 23  
 <211> 679  
 <212> DNA  
 <213> *Schizophyllum commune*

<220>  
 <223> (1-679) SC3 coding sequence

<220>  
 <223> (1-92) 1st cDNA

<220>  
 <223> (146-183) 2nd cDNA

<220>  
 <223> (240-317) 3rd cDNA

<220>  
 <223> (374-469) 4th cDNA

<220>  
 <223> (524-586) 5th cDNA

<220>  
 <223> (635-679) 6th cDNA

<400> 23  
 atgttcgccc gtctccccgt cgtgttcctc taagccttcg tcgcgttcgg cgccttcgctc 60  
 gctgccctcc caggtggcca cccgggcacg acgtacgtcg acctctcacc gtcctctaata 120  
 gtcttgctga tgaagccccg tatagcacgc cgccgggtac gacgacgggt acggtgacca 180  
 cggtagtagt ctttctcgcc gtgcgacgact cgaacgcatt ggctaatttt tgctcatagc 240  
 cgccttcgac gacgaccatc gccgcgggtg gcacgtgtac tacgggggtcg ctctcttgct 300  
 gcaaccagggt tcaatcggtc cgtacatcaa agcgggcacga ccaggcatct cagctgacgg 360  
 ccacatcgta caggcgagca gcagccctgt taccgccctc ctgggcctgc tcggcattgt 420  
 cctcagcgac ctcaacgttc tcgttggtat cagctgctct cccctcactg tgagatcttt 480  
 ttgttctactg tcccaattac tgcgcactga cagactttgc caggctcatc gtgtcggagg 540  
 cagcggtgtg tcggcgagca ccgtctgtg cgaaaacacc caattcgtat gtatactttc 600  
 catgctgtgc cctttctccg ctaatcatct gtagaacggg ctgatcaaca tcggttgac 660  
 ccccatcaac atcctctga 679

<210> 24  
 <211> 63  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PCR 5' primer

<400> 24  
 actacacgga ggagctcgac gacttcgagc agcccagct gcacgcaggg tggccacccg 60  
 ggc 63

<210> 25  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 25  
tcgtacggat cctcagagga tgttgatggg 30

<210> 26  
<211> 43  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 26  
ggaattccgc ggactgcgca tcatgaagtt cttcgccatc gcc 43

<210> 27  
<211> 80  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 27  
tgaattccat atgttaggta ccaccggggc ccatgccggt agaagtagaa gccccgggag 60  
caccgacggc ggtctggcac 80

<210> 28  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 28  
tgaattcggc acccaggctt gctcaagcgt c 31

<210> 29  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 29  
tgaattccat atgtcacagg cactgagagt agta 34

<210> 30  
<211> 48  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 30  
gaattcggta ccctcgtccc tcgcggtccc gccgaagtga acctgggtg 48

<210> 31  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 31  
tgaattccat atgctaaccc cgtttcatct ccag 34

<210> 32  
<211> 918  
<212> DNA  
<213> *Trichoderma reesei*

<220>  
<221> terminator  
<222> (1)..(918)  
<223> *T. reesei hfbI* terminator

<400> 32  
gatgcccgcc cggggtcagg gtgtgcccgt gagaaagccc acaaagtgtt gatgaggacc 60  
atttccggta ctgggaaagt tggctccacg tgtttgggca ggtttgggca agttgtgtag 120  
atattccatt cgtacgccat tcttattctc caatatttca gtacactttt cttcataaat 180  
caaaaagact gctattctct ttgtgacatg ccggaaggga acaattgtct ttggtctctg 240  
ttatttgcaa gtaggagtgg gagattcgcc ttagagaaag tagagaagct gtgcttgacc 300  
gtgggtgtgac tcgacgagga tggactgaga gtgttaggat taggtcgaac gttgaagtgt 360  
atacaggatc gtctggcaac ccacggatcc tatgacttga tgcaatgggt aagatgaatg 420  
acagtgtgaa aggaaaagga aatgtccgcc ttcagctgat atccacgcca atgatacagc 480  
gatatacctc caatatctgt gggaaacgaga catgacatat ttgtgggaac aacttcaaac 540  
agcgagccaa gacctcaata tgcacatcca aagccaaaca ttggcaagac gagagacagt 600  
cacattgtcg tcgaaaagatg gcatcgtagc caaatcatca gctctcatta tcgcctaaac 660  
cacagattgt ttgccgtccc ccaactccaa aacgttacta caaaagacat gggcgaatgc 720  
aaagacctga aagcaaacc tttttgcgac tcaattccct cctttgtcct cggaatgatg 780  
atccttcacc aagtaaaaga aaaagaagat tgagataata catgaaaagc acaacggaaa 840  
cgaaagaacc aggaaaagaa taaatctatc acgcaccttg tccccacact aaaagcaaca 900  
ggggggggtaa aatgaaat 918



<210> 33  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PCR 5' primer

<400> 33  
 gacctcgatg cccgcccggt gtcaag 26

<210> 34  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PCR 3' primer

<400> 34  
 gtcgacattt catatttacc ccctcg 26

<210> 35  
 <211> 1190  
 <212> DNA  
 <213> *Trichoderma reesei*

<220>  
 <221> promoter  
 <222> (1)..(1190)  
 <223> *T. reesei hfb2* promoter

<400> 35  
 ctcgagcagc tgaagcttgc atgcctgcat cctttgtgag cgactgcac cattttgcac 60  
 acactgccgt cgacgtctct cttccgacct tggccagctg gacaagcaac acaccaatga 120  
 cgctttgtat tattagagta tatgcaagtc tcaggactat cgactcaact ctaccaccg 180  
 aggacgatcg cggcacgata cgccctcgtt ctcatggcc caagcagacc aactgccct 240  
 ggagcaagat tcagcccaag ggagatggac ggcagggcac gccaggcccc caccaccaag 300  
 ccactccctt tggccaaatc agcttgcatt tcaagagaca tcgagctgtg ccttgaaatt 360  
 actaacaacc agggatggga aacgaagcct gcttttgga agacaacaat gagagagaga 420  
 gagagagggg gagagacaat gagtgccaca aacctggtag tgctccgcca atgcgtctga 480  
 aatgtcacat ccgagtcctg gggcctctgt gagaatgtcc agagtaatac gtgttttgcg 540  
 aatagtcctc tttcttgagg actggatacc tacgataccc tttttgagtt gatgcggtgc 600  
 tttcgaagta ttatctggag gatagaagac gtctaggtaa ctacacaaaa ggccataact 660  
 ttggggagta gcccaacgaa aggtaactcc tacggcctct tagagccgtc atagatccta 720  
 cagcctcttg gagccgtcat agatcacatc tgtgtagacc gacattctat gaataatcat 780  
 ctcatcatgg ccacatacta ctacatacgt gtctctgcct acctgacatg tagcagtggc 840  
 caagacacca aggccccagc atcaagcctc cctacctatc ccttccattg tacagcggca 900  
 gagagattgc gatgagccct ctccctacct acagacggct gacaatgtcc gtataccacc 960  
 agccaacgtg atgaaaacaa ggacatgagg aacagcctgc gagagctgga agatgaagag 1020  
 ggccagaaaa aaaagtataa agaagacctc gattcccgcc atccaacaat cttttccatc 1080  
 ctcatcagca cactcatcta caaccatcac cacattcact caactcctct ttctcaactc 1140  
 tccaaacaca aacattcttt gttgaatacc aaccatcacc acctttcaag 1190

<210> 36  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 36  
aagcttgcat gcctgcatcc 20

<210> 37  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 37  
ccatggtgaa aggtggtgat ggttgg 26

<210> 38  
<211> 13  
<212> PRT  
<213> *Trichoderma reesei*

<220>  
<223> vild type *T. reesei* EGI peptide linker

<400> 38  
Val Pro Arg Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly  
1 5 10

<210> 39  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: modified CBHII linker

<400> 39  
Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly  
1 5 10

<210> 40  
<211> 19  
<212> PRT  
<213> Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Met/Thrombin  
linker

&lt;400&gt; 40

Pro Gly Arg Pro Val Leu Thr Gly Pro Gly Met Gly Thr Ser Thr Ser  
1 5 10 15

Ala Gly Pro

&lt;210&gt; 41

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Met-containing linker

&lt;400&gt; 41

Pro Gly Ala Ser Thr Ser Thr Gly Met Gly Pro Gly Gly  
1 5 10

&lt;210&gt; 42

&lt;211&gt; 14

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: linker containing the thrombin  
cleavage site

&lt;400&gt; 42

Gly Thr Leu Val Pro Arg Gly Pro Ala Gly Val Asn Leu Val  
1 5 10